RESCUE AND PRESERVATION OF SAMPLE DATA FROM THE APOLLO MISSIONS TO THE MOON.

Nancy S. Todd[1], Ryan A. Zeigler[2], Cindy A. Evans[2], Kerstin Lehnert[3], [1] UTC Aerospace Systems/JETS Contract, NASA Johnson Space Center, Mail Code XI2, Houston, TX 77058; nancy.s.todd@nasa.gov. [2] NASA Johnson Space Center, 2101 NASA Rd 1, Mail Code XI2, Houston, TX 77058. [3]Lamont-Doherty Earth Observatory, Columbia University, 61 Rt 9W, Palisades, NY, 10964.

Introduction: Six Apollo missions landed on the Moon from 1969-72, returning to Earth 382 kg of lunar rock, soil, and core samples. These samples are among the best documented and preserved samples on Earth that have supported a robust research program for 45 years. From mission planning through sample collection, preliminary examination, and subsequent research, strict protocols and procedures are followed for handling and allocating Apollo subsamples, resulting in the production of vast amounts of documentation. Even today, hundreds of samples are allocated for research each year, building on the science foundation laid down by the early Apollo sample studies and combining new data from today's instrumentation, lunar remote sensing missions and lunar meteorites.

Much sample information is available to researchers at curator.jsc.nasa.gov. Decades of analyses on lunar samples are published in LPSC proceedings volumes and other peer-reviewed journals, and tabulated in lunar sample compendia entries. However, for much of the 1969-1995 period, the processing documentation, individual and consortia analyses, and unpublished results exist only in analog forms or primitive digital formats that are either inaccessible or at risk of being lost forever because critical data from early investigators remain unpublished.

Sample Documentation Process: The Apollo sample collection is unique amongst NASA's Astromaterials collections in that it is the only sample suite collected by astronauts, and the only suite collected with detailed geologic context. As such, the documentation process began before the samples were even collected. Properly documenting the sample history involves: (a) documenting the details of the sample's provenance, which helps with the interpretation of the local and regional geologic setting; (b) documenting the processing history of the sample, which allows studies separated by decades to be correlated; and (c) recording the analytical history of returned samples, which allows for the best use of returned samples. Thus, data collected can be characterized as belonging to one of three phases in the sample's lifecycle, as seen in Figure 1.

Existing Data Products. In the 45 years since the Apollo samples were collected, we have amassed large amounts of information of varying complexity. The geologic context of the samples was recorded in maps, images, and transcripts from the active Apollo missions.

Figure 1. Sample Documentation Phases

Sample Acquisition & Transport
- Collection place, time, and conditions
- Handling and storage during transport

Sample Curation & Processing
- Processing history
- Sample tracking and auditing

Sample Analysis
- All actions taken on the sample by the PI before they are returned to JSC
- All published and unpublished relevant analysis results from the sample analyses

This material has been subsequently digitized and made available to the public as PDF files but work still remains to correlate the available data to the individual Apollo samples in a comprehensive manner.

Sample processing of the Apollo collections at JSC has yielded a huge volume of information, from the day-to-day records of sample processing, to different types of documentation images (the initial whole rock images, processing images during subdivision of the samples, and thin section images). As is the case with most legacy data, these data exist primarily in the form of handwritten notes and annotated images leading to challenges in making this data publicly available to inform PIs about what studies about the Moon are feasible.

Lunar samples have been allocated >3000 times to PIs since 1969. These allocations have yielded well in excess of 5000 peer reviewed articles that contain geochemical or geophysical data about the lunar samples. Moreover, each of these studies has generated large amounts of very useful unpublished data that has not seen the light of day (materials that would often today go into electronic appendices). The closest thing we have to a compilation of this information is the Lunar Sample Compendium, spearheaded by Chuck Meyer [1] of JSC Astromaterials Curation, which summarizes the important geochemical results for most Apollo samples. However, the format of this compendium does not lend itself well to data discovery for researchers.

Data Management Challenges: In order to best protect and utilize the Apollo samples for future study, we need to address the issues associated with the management of sample data. Some of the challenges we face include:

Size and Complexity. The sheer size and complexity of the data products involved, each of which has different requirements, e.g., samples, maps, images, handwritten processing data, transcripts, published and unpublished analytical data, etc., not only requires large amounts of resources and time to process, but also requires resources with expertise in a variety of areas.

Age of the Data. There is a large number of incomplete or inaccessible data related to the age of the records, e.g., missing metadata, obsolete file types, illegible or incomplete processing notes, unpublished PI data. It is important to rescue and restore as much of these data as possible but it requires access to a rapidly dwindling set of experts who have the necessary knowledge to fill in the blanks.

Data Integrity. Data quality is critical in the successful recovery and restoration of existing data. However, much of the available data suffers from issues related to poor data control. Several issues have contributed greatly to the decreased integrity of data, including the use of hand-written notes, unstructured data formats (such as catch-all comment fields), and changes in the data collection processes. Changes in the state of technology, advances in sample processing and analysis techniques, and accumulated experience all lead to changes in the documentation process and in the data itself that are seldom addressed in legacy data retroactively. This leads to data that is inconsistent, difficult to import into databases, and ultimately inaccessible to the public. It is critical that we address existing issues of data quality by cleaning up and standardizing existing data. However, it is just as important that we implement data governance and quality control principles to preserve the integrity of the data in the future.

Technological Challenges. Technology greatly impacts the long term archival and accessibility needs of our data. For data that is born digital, this presents an even greater challenge. Safeguards must be put in place to ensure the survivability of the raw data despite the obsolescence of existing technologies.

Technology also affects data access and availability. For data to be made truly available to the public, we must implement delivery systems with often costly infrastructure requirements. We must also ensure that the data exists in formats that are usable and accessible by other systems outside Curation to ensure maximum searchability and access for users.

Data Recovery Initiatives: We have initiated several efforts to rescue some of the early Apollo data, including unpublished analytical data.

Lunar Image Scanning Project. This is a four-year LASER-funded project to digitize the original film negatives of Apollo Lunar Rock Sample photographs, which are the only existing record of the rocks as they were when received at JSC [2]. This effort resulted in the digitization of over 50K images, which are available to the public on the Curation Lunar Sample and Photo Catalog database. The raw images obtained from this effort are also archived in TIFF format and available through the PDS Imaging node.

Digitization of Analog Data Products. In house effort to digitize at-risk analog data products that document the Apollo samples, such as sample data packs and PI sample accountability forms (F75s). These records encompass over 400 feet of files kept in a vault in the Sample Control Center. Completed scanning of existing records produced the following: for data packs, ~60GB of data in 55K files; for F75s, ~90 GB in 46K files.

Further work is needed to rescan the data pack photos at a higher resolution and to render data that can be directly ingested into a database and made searchable.

MoonDB. The MoonDB project is a NASA-funded effort led by the Interdisciplinary Earth Data Alliance (IEDA), in conjunction with JSC curation. MoonDB is intended to be a quality-controlled data system that will preserve, digitize, and curate lunar geochemical, geochronological, and petrological data and their associated sample metadata and analytical metadata. For more information, see LPSC abstract #2738.

Other Initiatives. Other initiatives include micro-CT scanning of complex lunar samples to document their interior structure (e.g. clasts, vesicles); linking high-resolution scans of Apollo film products to samples; and new procedures for systematic high resolution photography of samples before additional processing, enabling detailed 3D reconstructions of the samples.

Conclusion: Properly managing and mantaining these comprehensive data records for future generations and making them readily available to Lunar researchers all over the world is of utmost importance in the preservation of the scientific legacy of the Apollo missions to the Moon. NASA's Astromaterials Curation is working on several initiatives aimed at preserving this legacy. All of these efforts will provide comprehensive access to Apollo samples and support better curation of the samples for decades to come.

References: [1] Lunar Sample Compendium, http://curator.jsc.nasa.gov/lunar/lsc. [2] Todd, N.S., Lofren,G.E., et al., *Apollo Lunar Sample Photograph Digitization Project Update*, LPSC 43, abstract #2860.